Form Automation in Call Center

Session 2022-2026

Prepared by

Abdul Rehman 2022-CS-57

Mudasir Ahmed 2022-CS-77

Submitted to:

Prof. Samyan Qayyum Wahla

Department of Computer Science

University of Engineering and Technology, Lahore

Contents

Abstract

The system designed to automate customer data entry. Admins can manage agent profiles, while agents handle customer details. Agents can input data manually or use audio recordings, which are transcribed and processed by a Roberta-based model for automatic form filling. The system extracts key information like name, address, phone number, and email from audio, reducing manual entry time, improving accuracy, and streamlining operations in industries where verbal interactions are converted into structured data.

1 Introduction

1.1 Description

The AI-powered call center management system is designed to streamline and optimize the handling of customer data, improving efficiency for both admins and agents. The system has two key user roles: the admin, responsible for managing agent profiles through Create, Read, Update, and Delete (CRUD) functionalities, and the agent, who manages customer details. Agents can view, update, or delete customer records they have previously added. A significant feature of this system is its ability to automate form completion by processing audio data. Instead of manually entering customer information, agents can record or upload an audio file containing the customers details. The system uses speech-to-text technology to convert the audio into a transcript and then sends the transcript to a Roberta-based Named Entity Recognition (NER) model. This model extracts key customer details such as name, address, phone number, email, and age from the transcript and automatically populates the respective fields in the form. Once the agent reviews and confirms the information, the form is saved, and a PDF version is automatically downloaded.

This AI-driven approach significantly reduces the time and effort agents typically spend on manual data entry, minimizing human errors and enhancing accuracy. The system is designed to address the challenges faced by industries where large volumes of forms are filled based on verbal interactions, providing an efficient solution to improve overall operational workflows.

1.2 Reason for Undertaking

The primary motivation behind this project is to reduce the manual workload for call center agents when processing customer data. Traditionally, agents spend a significant amount of time filling out forms based on verbal conversations with customers. This manual data entry process is not only time-consuming but also prone to errors.

By automating the extraction of essential customer details from audio recordings and autofilling forms, this system significantly improves efficiency. The AI-powered solution aims to enhance accuracy, save time, and minimize human errors in industries where form completion is a crucial task.

2 Functionalities

2.1 User Authentication

The system provides secure authentication for both admins and agents using JSON Web Tokens. Admins can perform CRUD operations on agent profiles, while agents can manage customer data within their own scope.

2.2 Speech-to-Text Conversion

Agents can either manually enter customer details or record audio. The system processes audio recordings and converts them into text using speech-to-text conversion. This feature simplifies data entry and enhances efficiency by eliminating the need for manual typing.

2.3 Data Labeling

Label Studio is used for labeling data to train the NER model. This tool helps in preparing labeled datasets, ensuring the model's accuracy in extracting the correct information from audio transcriptions.

2.4 Information Extraction (NER)

After the speech-to-text conversion, the transcribed text is sent to a pre-trained Roberta model for Named Entity Recognition (NER). The model extracts essential details such as the customers name, address, phone number, email, and age, making the process of filling out customer forms more accurate and faster.

2.5 Form Automation

Once the necessary details are extracted through NER, the system automatically fills out the customer details form. This reduces manual data entry time and minimizes errors. The form is then saved as a downloadable PDF, making it ready for use.

2.6 User Experience

The system provides a user-friendly interface where agents and admins can easily navigate through tasks such as managing customers, viewing data, and generating PDFs. The

integration of AI tools enhances the overall user experience by automating time-consuming tasks.

3 Technologies Used

The AI-Powered Call Center Management System uses various technologies to provide seamless data management and automated form filling. Below is a table summarizing the key technologies used in this project:

Technology	Description
MERN Stack	Full-stack development for the website, consisting of: **MongoDB:** Database for storing customer and agent data. **Express.js:** Backend framework for handling API requests. **React.js:** Frontend library for building user interfaces. **Node.js:** Server-side environment for executing JavaScript.
Roberta Model	Pre-trained Roberta model used for Named Entity Recognition (NER). It extracts customer details such as name, address, phone number, email, and age from transcribed audio data.
Multer	Middleware for handling file uploads, used to manage audio recordings.
FileSaver.js	JavaScript library for saving forms as PDF files.
Label Studio	Open-source tool used for annotating and labeling data for NER training.

Table 1: Technologies Used in AI-Powered Call Center Management System

4 AI Perspective(NLP)

The AI model used in this system is designed to perform Named Entity Recognition (NER), which is a critical task in Natural Language Processing (NLP) for identifying and classifying key pieces of information from text data.

4.1 Named Entity Recognition (NER)

Named Entity Recognition (NER) is a subtask of Information Extraction (IE) that focuses on identifying named entities within a text. These entities can include person name, location, phone, etc. In this system, the NER model is used to extract these critical

details from customer conversations. This enables agents to quickly capture and process relevant information from unstructured data.

4.2 Role of the AI Model in the System

The AI model plays a central role in transforming spoken customer data into structured and meaningful information. By converting unstructured data into structured formats, it reduces manual data entry and improves data accuracy. Specifically, in this system, the AI model:

- **Processes Audio Data:** When an agent records audio or uploads a file, the system first converts the audio into text (speech-to-text), and then the NER model processes this text to extract entities.

- **Auto-Fills Forms:** Based on the extracted entities, the AI model helps auto-populate fields such as customer name, phone number, email, etc., thus streamlining the data entry process for the agent.

- **Reduces Human Error:** By using AI, the system try to extract accurate and consistent details.

4.3 The Model Architecture: RoBERTa

The core of the systems AI is based on RoBERTa (A Robustly Optimized BERT Pre-training Approach), a variant of the well-known BERT (Bidirectional Encoder Representations from Transformers) model. RoBERTa is designed for sequence-to-sequence tasks like NER and is known for its ability to handle contextual language understanding. The model architecture is as follows:

- **Pretraining:** RoBERTa is pretrained on large text corpora using unsupervised learning, which allows it to learn rich representations of language. It becomes really good at understanding the meaning of words and how they are connected to each other. This ability makes it excellent at identifying important information.

- **Token Classification:** The model is fine-tuned to classify tokens (words or subwords) as different types of entities.

- **Transformer Layers:** It uses multiple transformer layers, which allow it to model complex relationships in the text. These layers provide deep contextual understand-

ing, which is necessary for accurately identifying entities based on their surrounding context.

- **Output Layer:**The model outputs a probability distribution for each token in the text. This distribution corresponds to the likelihood of each token belonging to a particular entity class (e.g., B-NAME, I-NAME).

4.4 Data Collection

Data for the call center system is mainly generated by code because there aren't any free datasets that accurately reflect real call center conversations. This means the system creates simulated conversations and customer interactions, which are then used to train the model. While this helps build a dataset, it can cause problems like overfitting. Since the data is artificial, it may not include all the different situations that happen in real-life calls. As a result, the model might perform well with the generated data but have trouble handling real-world data that it hasn't seen before.

4.5 Data Labeling

We have used Label Studio, an open-source data labeling tool, to label the data. Label Studio provided an intuitive interface to annotate entities such as names, emails, phone numbers, and addresses in text, allowing us to create high-quality labeled data for training our model.

4.6 Model Training and Fine-Tuning

We have fine-tuned it on a labeled dataset. This involves using annotated conversation data that contains both text and entity labels. The model learns to associate specific tokens with entity types based on the training data. Heres how fine-tuning works:

- **Supervised Learning:**The model is trained using supervised learning, where it learns from labeled examples. The labels correspond to the named entities in the text (e.g., B-NAME for names, I-EMAIL for email addresses).

- **Optimization:** The model is optimized using a loss function, which measures the difference between predicted labels and true labels. The optimizer (AdamW) adjusts the models parameters to minimize this loss.

5   Links

Form Automation Demo Video: YouTube Video
Model: NER roBERTa AI
Dataset: NER roBERTa AI Dataset